

One optional observation inflates α by $100/\sqrt{n}$ per cent

BY LUTZ MATTNER

Universität Trier

OptionalObservationsMetrika.V2.20090425.tex

For one-sample level α tests ψ_m based on independent observations X_1, \dots, X_m , we prove an asymptotic formula for the actual level of the test rejecting if at least one of the tests $\psi_n, \dots, \psi_{n+k}$ would reject. For $k = 1$ and usual tests at usual levels α , the result is approximately summarized by the title of this paper.

Our method of proof, relying on some second order asymptotic statistics as developed by Pfanzagl and Wefelmeyer, might also be useful for proper sequential analysis. A simple and elementary alternative proof is given for $k = 1$ in the special case of the Gauss test.

1. Main result and examples.

1.1. *Introduction and main result.* For a given one-sample testing problem and for every sample size m , let ψ_m be a test of level α , based on the m independent observations X_1, \dots, X_m . Suppose that initially n observation were planned, but that these do not lead to the desired rejection of the hypothesis. Then some experimenters might be tempted to collect up to k further observations X_{n+1}, \dots, X_{n+k} , calculating after each the test based on the accumulated observations, and to declare in effect a rejection of the hypothesis at level α if $\psi_m = 1$ for some $m \in \{n, \dots, n+k\}$. This would of course be wrong, but by how much? Surprisingly this question, known in the statistical literature at least since the publications of Feller [7] and Robbins [17], is usually not addressed in textbooks or treatises of statistics, see Subsection 1.2 below.

The title of the present paper gives a somewhat rough but easy to grasp answer for the simplest case of $k = 1$, approximately valid for common values of α and rather general one-sample tests based on asymptotically normal test statistics. Theorem 1.1 below gives a mathematically precise answer also for general k . We may summarize its statistical meaning as follows: Even an apparently slight amount of optional stopping will usually inflate the nominal

AMS 2000 subject classifications: Primary 62F05; secondary 62E20, 62L10.

Keywords and phrases: Abuse of statistical tests, level inflation, multiple testing, optional stopping, sequential test, statistics with stochastic expansion, teaching of statistics.

level of a test by a serious amount, such as by about 10 per cent for $n = 100$ and $k = 1$.

In our formulation of Theorem 1.1, we think of non-randomized tests ψ_m based on upper test statistics T_m with critical value zero, that is, $\psi_m = (T_m > 0)$, using the indicator notation (statement) $:= 1$ or 0 according to whether “statement” is true or false. Thinking only of tests actually exhausting a given level α , we essentially assume that this level is attained for at least one distribution from the hypothesis, simultaneously for all sufficiently large sample sizes. Theorem 1.1 refers to such a distribution, compare assumption (1) below, where the above qualifier “essentially” has been made precise as “ $+ o(1/\sqrt{n})$ ”. Unfortunately this assumption already excludes lattice cases like the binomial tests, for which any analogue of Theorem 1.1 would presumably look more complicated. Now the test rejecting if at least one of the tests $\psi_n, \dots, \psi_{n+k}$ would reject is $(\max_{m=n}^{n+k} T_m > 0)$, and hence, with respect to a given distribution of X_1 , its probability of rejecting is $\alpha_{n,k}$ as defined in (2) below. Our regularity assumptions (3)–(7) on the sequence (T_n) are similar to those imposed by Pfanzagl and Wefelmeyer in their well-known treatise of second order asymptotic statistics, see in particular [15, Section 10.3], on which our result is based. In Subsection 2.1 below we comment on some minor differences between these assumptions. Let Φ and $\varphi = \Phi'$ denote distribution function and density of the standard normal distribution $N_{0,1}$, and let us put

$$h(\alpha) := \frac{\varphi(\Phi^{-1}(1 - \alpha))}{\alpha\sqrt{2\pi}} \quad (\alpha \in]0, 1[)$$

We write $A_{\neq}^2 := \{(x, y) \in A^2 : x \neq y\}$ for any set A and $x_+ := x \vee 0 = \max\{x, 0\} = (-x)_-$ for $x \in \mathbb{R}$. Proofs of Theorem 1.1 and Lemmas 1.1 and 1.2 are given in Section 2, see Subsections 2.3, 2.4, and 2.5. We point out that Example 1.3.1 below contains an elementary direct proof, suitable for inclusion in standard statistics courses, of Theorem 1.1 in its simplest special case of the Gauss test with one optional observation, that is, $k = 1$.

THEOREM 1.1. *Let \mathcal{X} be a measurable space, $(X_n)_{n \in \mathbb{N}}$ a sequence of independent and identically distributed \mathcal{X} -valued random variables, $\alpha \in]0, 1[$, and $(t_n)_{n \in \mathbb{N}}$ a sequence of measurable functions $t_n : \mathcal{X}^n \rightarrow \mathbb{R}$ such that the random variables*

$$T_n := t_n(X_1, \dots, X_n) \quad (n \in \mathbb{N})$$

satisfy

$$(1) \quad \mathbb{P}(T_n > 0) = \alpha + o\left(\frac{1}{\sqrt{n}}\right) \quad (n \rightarrow \infty)$$

Let

$$(2) \quad \alpha_{n,k} := \mathbb{P}(\max_{m=n}^{n+k} T_m > 0), \quad \rho_{n,k} := \frac{\alpha_{n,k}}{\alpha} - 1 \quad (n, k \in \mathbb{N})$$

Assume that for $n \in \mathbb{N}$

$$(3) \quad T_n = \mu_0 + \frac{1}{\sqrt{n}} \sum_{i=1}^n f_0(X_i) + \frac{1}{\sqrt{n}} \left(\mu_1 + \frac{1}{\sqrt{n}} \sum_{i=1}^n f_1(X_i) + \frac{1}{2n} \sum_{(i,j) \in \{1, \dots, n\}^2_{\neq}} f_2(X_i, X_j) \right) + R_n$$

for some constants $\mu_0, \mu_1 \in \mathbb{R}$, measurable functions $f_0, f_1 : \mathcal{X} \rightarrow \mathbb{R}$ and $f_2 : \mathcal{X}^2 \rightarrow \mathbb{R}$, and a sequence $(R_n)_{n \in \mathbb{N}}$ of real-valued random variables with

$$(4) \quad \mathbb{E}f_0(X_1) = \mathbb{E}f_1(X_1) = \mathbb{E}(f_2(X_1, X_2)|X_1) = 0, \quad f_2(X_1, X_2) = f_2(X_2, X_1)$$

$$(5) \quad \mathbb{E}(f_0(X_1))^2 = 1, \quad \mathbb{E}|f_0(X_1)|^3 < \infty, \quad f_0(X_1) \text{ has a non-lattice law}$$

$$(6) \quad \mathbb{E}|f_1(X_1)|^{3/2} < \infty, \quad \mathbb{E}|f_2(X_1, X_2)|^{2+\delta} < \infty \text{ for some } \delta > 0$$

$$(7) \quad \text{For every } \varepsilon > 0: \quad \sup_{t \geq 1} t \mathbb{P}(|R_n| > \frac{t\varepsilon}{\sqrt{n}}) = o(\frac{1}{\sqrt{n}}) \quad (n \rightarrow \infty)$$

Then

$$(8) \quad \rho_{n,k} = \frac{h(\alpha)}{\sqrt{n}} \sqrt{2\pi} \sum_{\ell=1}^k \frac{1}{\ell} \mathbb{E} \left(\sum_{i=1}^{\ell} f_0(X_i) \right)_+ + o\left(\sqrt{\frac{k}{n}}\right) \quad \left(\frac{k}{n} \rightarrow 0\right)$$

and

$$(9) \quad \rho_{n,k} = 2h(\alpha) \sqrt{\frac{k}{n}} + o\left(\sqrt{\frac{k}{n}}\right) \quad \left(\frac{k}{n} \rightarrow 0, k \rightarrow \infty\right)$$

For common levels α , we have $h(\alpha) \approx 1$:

LEMMA 1.1. *The function h is strictly decreasing with the asymptotic behaviour*

$$(10) \quad h(\alpha) \sim \sqrt{\frac{1}{\pi} \log\left(\frac{1}{\alpha}\right)} \quad (\alpha \rightarrow 0)$$

and rounded values

α	0.05	0.025	0.01	0.005	0.001	0.0005
$h(\alpha)$	0.82	0.93	1.06	1.15	1.34	1.42

Taking $k = 1$ in (8), we get

$$(11) \quad \rho_{n,1} \sim \frac{h(\alpha)}{\sqrt{n}} \sqrt{2\pi} \mathbb{E}(f_0(X_1))_+ \quad (n \rightarrow \infty)$$

So, assuming $h(\alpha) \approx 1$, the claim in the title of this paper approximately results when $\sqrt{2\pi} \mathbb{E}(f_0(X_1))_+ \approx 1$ and n is sufficiently large. For the Gauss and t -tests in Examples 1.3.1 and 1.3.3 below, we have $\sqrt{2\pi} \mathbb{E}(f_0(X_1))_+ = 1$ exactly. In these two cases, from (15) below, one optional observation inflates α by $h(\alpha)100/\sqrt{n}$ per cent, two optional observations inflate α by $h(\alpha)171/\sqrt{n}$ per cent, etc. For general examples we note, using (4), (5) and $\mathbb{E}|Y| \leq (\mathbb{E}Y^2)^{1/2}$, that $\sqrt{2\pi} \mathbb{E}(f_0(X_1))_+$ can be any strictly positive number $\leq \sqrt{2\pi} \frac{1}{2} (\mathbb{E}(f_0(X_1))^2)^{1/2} = \sqrt{\pi/2}$, so the accuracy of the claim in the title depends on $\sqrt{2\pi} \mathbb{E}(f_0(X_1))_+$ being not too far from its value under $f_0(X_1) \sim N_{0,1}$. In the exponential Example 1.3.2, we have $\sqrt{2\pi} \mathbb{E}(f_0(X_1))_+ = \sqrt{2\pi}/e = 0.92$, so that in this case one optional observation inflates α by merely $h(\alpha)92/\sqrt{n}$ per cent.

Many test sequences (ψ_n) in the literature can be written in the form $\psi_n = (T_n > 0)$ with (T_n) admitting an expansion as in Theorem 1.1. This is in particular true, under appropriate regularity conditions, for one-sided tests based on one-dimensional components of minimum contrast estimators, see [15, pp. 395-396, Theorem 11.3.4] for a precise statement and references. In our examples in Subsection 1.3 below we can easily check all assumptions rather directly.

We have to note here that our assumption (7) on the sequence of remainders (R_n) is slightly stronger than Pfanzagl and Wefelmeyer's

$$(12) \quad \text{For every } \varepsilon > 0: \quad \mathbb{P}(|R_n| > \varepsilon/\sqrt{n}) = o(1/\sqrt{n}) \quad (n \rightarrow \infty)$$

Condition (7) appears to be just about what is needed in the proof of our crucial Lemma 2.1 below, since we allow k to be unbounded, see (21) below. For bounded k , assumption (12) would suffice. Condition (7) should be easy to establish in any reasonable case, and we do this in Example 1.3.3 below by using the following simple fact.

LEMMA 1.2. *Let $(R_n)_{n \in \mathbb{N}}$ be a sequence of real-valued random variables such that for some $p \in [1, \infty[$ and $n_0 \in \mathbb{N}$ the random variables*

$$(13) \quad Y_n := |n^{\frac{1+p}{2p}} R_n|^p$$

with $n \geq n_0$ are uniformly integrable. Then (7) holds.

For a discussion of further minor differences between Pfanzagl and Wefelmeyer's and our assumptions on (T_n) see Subsection 2.1 below.

1.2. *Various remarks.* Reading this subsection is not logically necessary for understanding the rest of this paper.

Under the assumptions of Theorem 1.1, one can easily show that $T_n - \mu_0 - (1/\sqrt{n}) \sum_{i=1}^n f_0(X_i)$ converges to zero in probability, so that $T_n - \mu_0$ converges in law to $N_{0,1}$, and hence, in view of (1), we must have

$$(14) \quad \mu_0 = -\Phi^{-1}(1 - \alpha)$$

The result of [15, Corollary 10.3.8], on which our proof of Theorem 1.1 will be based below, further includes a formula for μ_1 in terms of α, f_0, f_1, f_2 and the law of X_1 .

The expectations occurring in formula (8) can be computed explicitly in some cases, see in particular Example 1.3.2 below and, more generally, [5]. We always have $\mathbb{E}(\sum_{i=1}^{\ell} f_0(X_i))_+ \sim \sqrt{\frac{\ell}{2\pi}}$ for $\ell \rightarrow \infty$, see the end of the proof of Theorem 1.1, and $\mathbb{E}(\sum_{i=1}^{\ell} f_0(X_i))_+ \geq \sqrt{\frac{\ell}{2}} \mathbb{E}(f_0(X_1))_+$ by [12, Corollary 1.3].

Relation (9) becomes false if the condition " $k/n \rightarrow 0$ " is replaced by " k/n bounded", since for k/n constant and sufficiently large a contradiction to $\alpha_{n,k} \leq 1$ would result.

As mentioned above, the problem of level inflation due to optional stopping is usually not addressed in textbooks or treatises of statistics. It was raised, perhaps for the first time in the literature, by Feller in 1940 in connection with apparently ill-conducted experiments concerning "extra-sensory perception", see [7, pp. 286-294] and references therein. Robbins [17, pp. 534-535] posed the problem of evaluating or bounding what we have called $\alpha_{n,k}$, and stated without proof a bound in the case of the Gauss test. We are not aware of a continuation of that part of Robbins' work. Diaconis [4] comments critically on Feller's paper, but not on the particular point of optional stopping. Among books known to the present author, Pfanzagl's [14, p. 127]) is unique in stressing and demonstrating the problem, albeit only by a simulation, and unfortunately obscured by the additional deliberate mistake of choosing between two valid test for each sample size. To our surprise, we did not find any statistical textbook treating the problem more systematically.

Theorem 1.1 can be read as addressing an improper sequential analysis. Its technical basis however, namely the consideration of statistics with stochastic expansion, Pfanzagl and Wefelmeyer's result on their asymptotic distributions, and the crucial Lemma 2.1 below, might be useful for proper sequential analysis as well. For example, under the assumptions of Theorem 1.1 but with

condition (1) omitted, we can generalize (8) to a computation of the asymptotic distribution of $\max_{m=n}^{n+k} T_m$ up to an error $o(\sqrt{k/n})$ for $k/n \rightarrow 0$, by using [15, Proposition 10.3.1] rather than [15, Corollary 10.3.8] in a modification of the present Proof 2.3.

1.3. *Examples.* In each case below, let $\alpha \in]0, 1[$ be the level of the tests considered.

1.3.1. *The Gauss test.* The Gauss test for testing $\mu \leq \mu_0$ based on i.i.d. normal X_1, \dots, X_n with unknown mean $\mu \in \mathbb{R}$ and known standard deviation $\sigma_0 \in]0, \infty[$ rejects iff

$$T_n := -\Phi^{-1}(1 - \alpha) + \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{X_i - \mu_0}{\sigma_0} > 0$$

Hence Theorem 1.1 is applicable, with $X_i \sim N_{\mu_0, \sigma_0^2}$, $f_0(x) := (x - \mu_0)/\sigma_0$, and vanishing μ_1 , f_1 , f_2 , and R_n , and (8) reads

$$(15) \quad \rho_{n,k} = \frac{h(\alpha)}{\sqrt{n}} \sum_{\ell=1}^k \frac{1}{\sqrt{\ell}} + o\left(\sqrt{\frac{k}{n}}\right) \quad \left(\frac{k}{n} \rightarrow 0\right)$$

Here is the elementary proof of (15) for the simplest case of $k = 1$ promised immediately before the statement of Theorem 1.1: With $Y_i := (X_i - \mu_0)/\sigma_0$, $Z_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i$ and $z := \Phi^{-1}(1 - \alpha)$, we have

$$\begin{aligned} \alpha_{n,1} - \alpha &= \mathbb{P}(Z_n > z \text{ or } Z_{n+1} > z) - \mathbb{P}(Z_n > z) \\ &= \mathbb{P}(Z_n \leq z, Z_{n+1} > z) \\ &= \mathbb{P}(Z_n \leq z, Y_{n+1} > \sqrt{n+1}z - \sqrt{n}Z_n) \\ &= \int_{-\infty}^z \left(1 - \Phi(\sqrt{n+1}z - \sqrt{n}t)\right) \varphi(t) dt \end{aligned}$$

since Z_n and Y_{n+1} are independent and $N_{0,1}$ -distributed. Hence, using the change of variables $t \mapsto z - \frac{t}{\sqrt{n}}$, we get

$$\begin{aligned} \sqrt{n} (\alpha_{n,1} - \alpha) &= \int_0^\infty \left(1 - \Phi((\sqrt{n+1} - \sqrt{n})z + t)\right) \varphi\left(z - \frac{t}{\sqrt{n}}\right) dt \\ &\xrightarrow{(n \rightarrow \infty)} \int_0^\infty (1 - \Phi(t)) \varphi(z) dt \\ &= \varphi(z) \int_0^\infty t \varphi(t) dt \\ &= \frac{\varphi(z)}{\sqrt{2\pi}} \end{aligned}$$

by dominated convergence with the integrands dominated by the function $t \mapsto (1 - \Phi(t - z_-))\varphi(0)$.

1.3.2. *Testing an exponential mean.* The usual optimal test for $\lambda \geq \lambda_0$ based on i.i.d. exponential X_1, \dots, X_n with density $]0, \infty[\ni x \mapsto \lambda e^{-\lambda x}$ with $\lambda \in]0, \infty[$ unknown rejects for large values of $\sum_{i=1}^n X_i$, namely iff

$$T_n := -F_{P_n}^{-1}(1 - \alpha) + \frac{1}{\sqrt{n}} \sum_{i=1}^n (\lambda_0 X_i - 1) > 0$$

where P_n denotes the law of the standardization of $\sum_{i=1}^n X_i$ under $\lambda = \lambda_0$ and $F_{P_n}^{-1}$ the corresponding quantile function. Since P_n admits an Edgeworth expansion with remainder $o(1/\sqrt{n})$, see e.g. [13, p. 174, Theorem 5.22], we have

$$T_n = -\Phi^{-1}(1 - \alpha) + \frac{1}{\sqrt{n}} \sum_{i=1}^n f_0(X_i) + \frac{\mu_1}{\sqrt{n}} + R_n$$

where $f_0(x) = \lambda_0 x - 1$, $\mu_1 \in \mathbb{R}$ depends only on α , and where R_n is deterministic and $o(1/\sqrt{n})$ for $n \rightarrow \infty$. Hence the assumptions of Theorem 1.1 are fulfilled. With $\gamma_a(x) := (\Gamma(a))^{-1} x^{a-1} e^{-x}$ we have $\int_x^\infty (t - a) \gamma_a(t) dt = a \gamma_{a+1}(x)$ for $a, x \in]0, \infty[$, by differentiation with respect to x and considering $x \rightarrow \infty$, so that

$$\mathbb{E} \left(\sum_{i=1}^{\ell} f_0(X_i) \right)_+ = \int_{\ell}^{\infty} (t - \ell) \gamma_{\ell}(t) dt = \left(\frac{\ell}{e} \right)^{\ell} \frac{1}{(\ell - 1)!}$$

and accordingly (8) reads

$$(16) \quad \rho_{n,k} = \frac{h(\alpha) \sqrt{2\pi}}{\sqrt{n} e} \sum_{\ell=1}^k \left(\frac{\ell}{e} \right)^{\ell-1} \frac{1}{(\ell - 1)!} + o \left(\sqrt{\frac{k}{n}} \right) \quad \left(\frac{k}{n} \rightarrow 0 \right)$$

1.3.3. *The t-test.* The t -test for $\mu \leq \mu_0$ based on i.i.d. normal X_1, \dots, X_n with unknown mean $\mu \in \mathbb{R}$ and unknown standard deviation $\sigma \in]0, \infty[$ rejects for $n \geq 2$ iff

$$T_n := \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i - c_n \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2} > 0$$

with $Y_i := (X_i - \mu_0)/\sigma_0$ with $\sigma_0 \in]0, \infty[$ arbitrary, $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$, and c_n denoting the $(1 - \alpha)$ -quantile of the t -distribution with $n - 1$ degrees of freedom. Assuming $X_i \sim N_{\mu_0, \sigma_0^2}$ from now on, the Y_i are standard normal. We write $\bar{Y}_n^2 := \frac{1}{n} \sum_{i=1}^n Y_i^2$ and $S_n^2 := \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2 = \bar{Y}_n^2 - \bar{Y}_n^2$ and use $c_n = \Phi^{-1}(1 - \alpha) + O(\frac{1}{n})$ e.g. from [11, p. 461, (11.75)], $\sqrt{\frac{n}{n-1}} = 1 + O(\frac{1}{n})$ for

$n \in \mathbb{N}$ with $n \geq 2$, and $\sqrt{x} = 1 + (x - 1)/2 + O((x - 1)^2)$ for $x \in [0, \infty[$, to get

$$\begin{aligned}
T_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i - c_n \sqrt{\frac{n}{n-1}} \sqrt{S_n^2} \\
&= \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i \\
&\quad - \left(\Phi^{-1}(1 - \alpha) + O\left(\frac{1}{n}\right) \right) \left(1 + \frac{1}{2}(\overline{Y_n^2} - 1) - \frac{1}{2}\overline{Y_n^2} + O((S_n^2 - 1)^2) \right) \\
&= -\Phi^{-1}(1 - \alpha) + \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i - \frac{\Phi^{-1}(1 - \alpha)}{2n} \sum_{i=1}^n (Y_i^2 - 1) + R_n
\end{aligned}$$

where the sequence (R_n) satisfies (7) by Lemma 1.2 with $p = 1$, since nR_n is a linear combination with bounded coefficients of the four random variables 1 , $\frac{1}{n} \sum_{i=1}^n (Y_i^2 - 1)$, $n\overline{Y_n^2}$, and $n(S_n^2 - 1)^2$, which are uniformly integrable, as may be verified by checking that their second moments are bounded. Hence the assumptions of Theorem 1.1 are fulfilled, with $f_0(x) = (x - \mu_0)/\sigma_0$, $\mu_1 = 0$, $f_1(x) = -\Phi^{-1}(1 - \alpha)((x - \mu_0)/\sigma_0)^2 - 1$, and $f_2 = 0$, and we get the same asymptotic formula (15) as in the Gauss case.

2. Auxiliary results and proofs. In this section we use Pfanzagl and Wefelmeyer's [15, p. 16] $\varepsilon_{\mathbb{P}}$ -notation: For real-valued random variables X_n on probability spaces $(\Omega_n, \mathcal{A}_n, \mathbb{P}_n)$ and numbers $\delta_n > 0$, we write

$$\begin{aligned}
X_n &= \varepsilon_{\mathbb{P}_n}(\delta_n) \quad (n \rightarrow \infty) \\
&: \iff \forall \varepsilon > 0 \quad \mathbb{P}_n(|X_n| \geq \varepsilon) = o(\delta_n) \quad (n \rightarrow \infty)
\end{aligned}$$

Here n can belong to any index set if “ $n \rightarrow \infty$ ” is replaced by the specification of some appropriate passage to the limit, formally by a filter or a net. In our case the index is actually $(n, k) \in \mathbb{N}^2$, but $\mathbb{P}_{n,k}$ is for notational convenience chosen to be independent of (n, k) , say an infinite product measure, so that $\varepsilon_{\mathbb{P}_{n,k}}$ becomes $\varepsilon_{\mathbb{P}}$. The three successively more specialized passages to the limit we use are “ $n \rightarrow \infty$ ”, “ $k/n \rightarrow 0$ ”, and “ $k/n \rightarrow 0, k \rightarrow \infty$ ”.

We begin with a comparison of our versus Pfanzagl and Wefelmeyer's assumptions on the stochastic expansion (3), then state and prove the crucial Lemma 2.1, and conclude by proving Theorem 1.1 and Lemmas 1.1 and 1.2.

2.1. Discussion of the assumptions (3), (5), and (7). Our assumptions on the sequence (T_n) differ in three respects from those of [15, p. 343, Corollary 10.3.8, $S_n = \mu(P) + T_n/\sqrt{n}$, the case $g_1 = \overline{g}_1 = 0$] used in the proof of Theorem 1.1 below.

First, to simplify the notation, we have added the normalizing assumption $\mathbb{E}(f_0(X_1))^2 = 1$.

Second, as already discussed above, we have (7) instead of Pfanzagl and Wefelmeyer's " $\sqrt{n} R_n = \varepsilon_{\mathbb{P}}(1/\sqrt{n})$ ", that is, (12).

Third, instead of our $U_n := \frac{1}{2n} \sum_{(i,j) \in \{1, \dots, n\}^2} f_2(X_i, X_j)$ in the stochastic expansion of T_n , in [15] we have $V_n := \frac{1}{2n} \sum_{(i,j) \in \{1, \dots, n\}^2} f_2(X_i, X_j)$ and the additional assumption $\mathbb{E}(f_2(X_1, X_1))^{3/2} < \infty$. Here our version is slightly more general, since, under the moment condition just stated, we have $V_n = U_n + \frac{1}{2} \mathbb{E} f_2(X_1, X_1) + \sqrt{n} R_n$, where the present $R_n := n^{-3/2} \sum_{i=1}^n \xi_i$ with $\xi_i := \frac{1}{2}(f_2(X_i, X_i) - \mathbb{E} f_2(X_1, X_1))$ also satisfies (7), as follows via Lemma 1.2 from the fact that for $p := 3/2$ the random variables Y_n defined by (13) are given by $Y_n = |n^{-2/3} \sum_{i=1}^n \xi_i|^{3/2}$ and hence are uniformly integrable by the Theorem of Pyke and Root [16]. Hence, even under our more stringent condition (7) on the remainders, we may in the expansion from [15] simultaneously replace V_n by U_n and μ_1 by $\mu_1 + \frac{1}{2} \mathbb{E} f_2(X_1, X_1)$. Moreover, [15, Proposition 10.3.1 and Corollary 10.3.8] remain true with U_n in place of V_n even if the assumption $\mathbb{E}(f_2(X_1, X_1))^{3/2} < \infty$ is omitted, since the latter is used in [15] only to replace V_n by U_n in the proof of [15, Proposition 10.3.1] in order to prepare for the application of the result of Bickel, Götze and van Zwet [2, Theorem 1.2] and Götze [8, Theorem 1.14] which refers to U -statistics rather than V -statistics.

Finally let us note that our non-latticeness assumption in (5) is the same as the one imposed in [15] using the confusing term "strongly non-lattice" necessary only for multivariate statistics T_n , see [1, pp. 207, 221, and 226]

2.2. The main lemma. The following Lemma 2.1 is the crucial first step in our proof of Theorem 1.1 in Subsection 2.3 below.

LEMMA 2.1. *Let $(T_n)_{n \in \mathbb{N}}$ be a sequence of real-valued random variables such that for $n \in \mathbb{N}$ we have (3) for some constants $\mu_0, \mu_1 \in \mathbb{R}$, a measurable space \mathcal{X} , a sequence $(X_n)_{n \in \mathbb{N}}$ of independent and identically distributed \mathcal{X} -valued random variables, measurable functions $f_0, f_1 : \mathcal{X} \rightarrow \mathbb{R}$ and $f_2 : \mathcal{X}^2 \rightarrow \mathbb{R}$, and a sequence $(R_n)_{n \in \mathbb{N}}$ of real-valued random variables with (4),*

$$(17) \quad \mathbb{E}|f_0(X_1)|^2 < \infty, \quad \mathbb{E}|f_2(X_1, X_2)|^{\frac{3}{2}+\delta} < \infty \text{ for some } \delta > 0$$

and (7). Then

$$(18) \quad \max_{m=n}^{n+k} \left| T_m - \left(T_n + \frac{1}{\sqrt{n}} \sum_{i=n+1}^m f_0(X_i) \right) \right| = \sqrt{\frac{k}{n}} \varepsilon_{\mathbb{P}} \left(\sqrt{\frac{k}{n}} \right) \quad \left(\frac{k}{n} \rightarrow 0 \right)$$

PROOF. Let $\varepsilon > 0$. We have for $n, k \in \mathbb{N}$

$$\begin{aligned}
U &:= \text{L.H.S.}(18) \\
&\leq \max_{m=n}^{n+k} \left| \left(\frac{1}{\sqrt{m}} - \frac{1}{\sqrt{n}} \right) \sum_{i=1}^m f_0(X_i) \right| + \max_{m=n}^{n+k} \left| \left(\frac{1}{\sqrt{m}} - \frac{1}{\sqrt{n}} \right) \mu_1 \right| \\
&\quad + \max_{m=n}^{n+k} \left| \frac{1}{m} \sum_{i=1}^m f_1(X_i) - \frac{1}{n} \sum_{i=1}^n f_1(X_i) \right| \\
&\quad + \max_{m=n}^{n+k} \left| \frac{1}{m^{3/2}} \sum_{j=2}^m \sum_{i=1}^{j-1} f_2(X_i, X_j) - \frac{1}{n^{3/2}} \sum_{j=2}^n \sum_{i=1}^{j-1} f_2(X_i, X_j) \right| \\
&\quad + \max_{m=n}^{n+k} |R_m - R_n| \\
&=: U_1 + U_2 + U_3 + U_4 + U_5
\end{aligned}$$

For $\alpha \in \mathbb{R}$, the elementary inequality

$$(1-x)^\alpha \geq 1 - (\alpha \vee 1) x \quad (x \in [0, 1])$$

applied to $x = k/(n+k)$ yields

$$(19) \quad \frac{1}{n^\alpha} - \frac{1}{(n+k)^\alpha} = \frac{1 - (1 - \frac{k}{n+k})^\alpha}{n^\alpha} \leq \frac{(\alpha \vee 1) k}{n^\alpha (n+k)} \leq \frac{(\alpha \vee 1) k}{n^{1+\alpha}}$$

Investigation of U_1 : By (19) with $\alpha = 1/2$,

$$U_1 \leq \frac{k}{\sqrt{n}(n+k)} \max_{m=n}^{n+k} \left| \sum_{i=1}^m f_0(X_i) \right|$$

and by Kolmogorov's inequality, see [6, p. 61],

$$\mathbb{P}(U_1 > \varepsilon \sqrt{\frac{k}{n}}) \leq \varepsilon^{-2} \mathbb{E}(f_0(X_1))^2 \frac{k}{n+k} = o\left(\sqrt{\frac{k}{n}}\right) \quad \left(\frac{k}{n} \rightarrow 0\right)$$

Investigation of U_2 : Again by (19) with $\alpha = 1/2$, we get

$$U_2 \leq \frac{|\mu_1|k}{\sqrt{n}(n+k)} = o\left(\sqrt{\frac{k}{n}}\right) \quad (n \rightarrow \infty)$$

Investigation of U_3 : By (19) with $\alpha = 1$, we get

$$\begin{aligned}
U_3 &= \max_{m=n}^{n+k} \left| \left(\frac{1}{m} - \frac{1}{n} \right) \sum_{i=1}^n f_1(X_i) + \frac{1}{m} \sum_{i=n+1}^m f_1(X_i) \right| \\
&\leq \frac{k}{n^2} \left| \sum_{i=1}^n f_1(X_i) \right| + \frac{1}{n} \max_{m=n}^{n+k} \left| \sum_{i=n+1}^m f_1(X_i) \right| \\
&=: U_{3,1} + U_{3,2}
\end{aligned}$$

By Markov's inequality and the L^1 -law of large numbers, see [6, p. 337],

$$\mathbb{P}\left(|U_{3,1}| > \varepsilon \sqrt{\frac{k}{n}}\right) \leq \frac{1}{\varepsilon} \sqrt{\frac{k}{n}} \mathbb{E}\left|\frac{1}{n} \sum_{i=1}^n f_1(X_i)\right| = o\left(\sqrt{\frac{k}{n}}\right) \quad (n \rightarrow \infty)$$

By Doob's inequality applied to the submartingale $(|\sum_{i=1}^\ell f_1(X_i)| : \ell \in \{0, \dots, k\})$, see [6, p. 247], and recalling our indicator notation (statement) := 1 or 0 according to whether “statement” is true or false,

$$\begin{aligned} \mathbb{P}\left(|U_{3,2}| > \varepsilon \sqrt{\frac{k}{n}}\right) &= \mathbb{P}\left(\max_{\ell=1}^k \left|\sum_{i=1}^\ell f_1(X_i)\right| > \varepsilon \sqrt{nk}\right) \\ (20) \quad &\leq \frac{1}{\varepsilon} \sqrt{\frac{k}{n}} \mathbb{E}\left(\left|\frac{1}{k} \sum_{i=1}^k f_1(X_i)\right| \cdot \left(\max_{\ell=1}^k \sum_{i=1}^\ell f_1(X_i) > \varepsilon \sqrt{nk}\right)\right) \\ &= o\left(\sqrt{\frac{k}{n}}\right) \quad (n \rightarrow \infty) \end{aligned}$$

where for the last step, given $\delta > 0$, we choose k_0 according to the L^1 -law of large numbers such that $\mathbb{E}\left|\frac{1}{k} \sum_{i=1}^k f_1(X_i)\right| < \delta/\varepsilon$ for $k > k_0$, and then n_0 such that for $k \leq k_0$ and $n \geq n_0$ the expectation in line (20) is $< \delta/\varepsilon$.

Investigation of U_4 : By (19) with $\alpha = 3/2$, we get

$$\begin{aligned} U_4 &= \max_{m=n}^{n+k} \left| \left(\frac{1}{m^{3/2}} - \frac{1}{n^{3/2}} \right) \sum_{j=2}^n \sum_{i=1}^{j-1} f_2(X_i, X_j) + \frac{1}{m^{3/2}} \sum_{j=n+1}^m \sum_{i=1}^{j-1} f_2(X_i, X_j) \right| \\ &\leq \frac{3k}{2n^{3/2}(n+k)} \left| \sum_{j=2}^n \sum_{i=1}^{j-1} f_2(X_i, X_j) \right| + \frac{1}{n^{3/2}} \max_{m=n}^{n+k} \left| \sum_{j=n+1}^m \sum_{i=1}^{j-1} f_2(X_i, X_j) \right| \\ &=: U_{4,1} + U_{4,2} \end{aligned}$$

Let $p := (\frac{3}{2} + \delta) \wedge 2$ for some δ from (17). With $c_1 := 2 \left(\frac{3}{2\varepsilon}\right)^p \mathbb{E}|f_2(x_1, X_2)|^p < \infty$, Markov's inequality and inequality (22) from Lemma 2.2 below applied to $f_{ij} := f_2$ yield

$$\begin{aligned} \mathbb{P}\left(|U_{4,1}| > \varepsilon \sqrt{\frac{k}{n}}\right) &\leq \left(\frac{3}{2\varepsilon}\right)^p k^{p/2} n^{-p} (n+k)^{-p} \mathbb{E}\left|\sum_{j=2}^n \sum_{i=1}^{j-1} f_2(X_i, X_j)\right|^p \\ &\leq c_1 k^{p/2} n^{2-p} (n+k)^{-p} \\ &= c_1 \sqrt{\frac{k}{n}} k^{(p-1)/2} n^{5/2-p} (n+k)^{-p} \\ &\leq c_1 \sqrt{\frac{k}{n}} n^{2-3p/2} \quad [\text{by } k \leq n+k \text{ and } n+k \geq n] \\ &= o\left(\sqrt{\frac{k}{n}}\right) \quad (n \rightarrow \infty) \end{aligned}$$

since $p > 4/3$. To bound $U_{4,2}$, we again use Lemma 2.2, but with $n+k$ in place of n and with $f_{ij} := f_2$ for $j > n$ and $f_{ij} := 0$ for $j \leq n$, to see that

$$M_m := \sum_{j=n+1}^m \sum_{i=1}^{j-1} f_2(X_i, X_j) \quad (m \in \{n+1, \dots, n+k\})$$

defines a martingal. Hence Doob's inequality, (22), $c_2 := 4\varepsilon^{-p}\mathbb{E}|f_2(x_1, X_2)|^p < \infty$, and $p > 3/2$ yield

$$\begin{aligned} \mathbb{P}\left(|U_{4,2}| > \varepsilon\sqrt{\frac{k}{n}}\right) &\leq \mathbb{P}\left(\frac{1}{n^{3/2}} \max_{m=n+1}^{n+k} |M_m| > \varepsilon\sqrt{\frac{k}{n}}\right) \\ &\leq (\varepsilon n\sqrt{k})^{-p} \mathbb{E}|M_{n+k}|^p \\ &\leq (\varepsilon n\sqrt{k})^{-p} 4 \sum_{j=n+1}^{n+k} \sum_{i=1}^{j-1} \mathbb{E}|f_2(X_i, X_j)|^p \\ &\leq c_2 \cdot (n\sqrt{k})^{-p} k(n+k) \\ &= c_2 \sqrt{\frac{k}{n}} \left(k^{\frac{1-p}{2}} n^{\frac{3}{2}-p} + k^{\frac{3-p}{2}} n^{\frac{1}{2}-p} \right) \\ &\leq c_2 \sqrt{\frac{k}{n}} \left(n^{\frac{3}{2}-p} + \frac{k^{\frac{3}{4}}}{n} \right) \\ &= o\left(\sqrt{\frac{k}{n}}\right) \quad \left(\frac{k}{n} \rightarrow 0\right) \end{aligned}$$

Investigation of U_5 : Using (7) with $t = \sqrt{k}$, we get

$$\begin{aligned} (21) \quad \mathbb{P}\left(|U_5| > 2\varepsilon\sqrt{\frac{k}{n}}\right) &\leq \sum_{m=n}^{n+k} \mathbb{P}\left(|R_m| > \varepsilon\sqrt{\frac{k}{m}}\right) \\ &= \frac{1}{\sqrt{k}} \sum_{m=n}^{n+k} \sqrt{k} \mathbb{P}\left(|R_m| > \varepsilon\sqrt{\frac{k}{m}}\right) \\ &= \frac{k+1}{\sqrt{k}} o\left(\frac{1}{\sqrt{n}}\right) \\ &= o\left(\sqrt{\frac{k}{n}}\right) \quad (n \rightarrow \infty) \end{aligned}$$

Combining the results for U_1, \dots, U_5 , we get

$$P(|U| > 8\varepsilon\sqrt{k/n}) = o\left(\sqrt{\frac{k}{n}}\right) \quad \left(\frac{k}{n} \rightarrow 0\right)$$

□

The following lemma, which we have just used above when handling U_4 , is in principle well known, see for example Koroljuk and Borovskich' book [10, p. 72, Theorem 2.1.3, the case $r = c = 2$] for the special case where the f_{ij} are symmetric and independent of (i, j) .

LEMMA 2.2. *Let X_1, \dots, X_n be independent \mathcal{X} -valued random variables and let $f_{ij} : \mathcal{X}^2 \rightarrow \mathbb{R}$ be measurable with $\mathbb{E}|f_{ij}(X_i, X_j)| < \infty$ and with $\mathbb{E}(f_{ij}(X_i, X_j)|X_i) = 0$ for $1 \leq i < j \leq n$. Then*

$$M_m := \sum_{j=2}^m \sum_{i=1}^{j-1} f_{ij}(X_i, X_j) \quad (m \in \{2, \dots, n\})$$

defines a martingale, and for $p \in [1, 2]$ we have

$$(22) \quad \mathbb{E}|M_n|^p \leq 4 \sum_{j=2}^n \sum_{i=1}^{j-1} \mathbb{E}|f_{ij}(X_i, X_j)|^p$$

PROOF. Clearly $(M_m : m \in \{2, \dots, n\})$ is a martingale with respect to the σ -algebras $\sigma(X_1, \dots, X_m)$, and so is $(\sum_{i=1}^m f_{ij}(X_i, X_j) : m \in \{2, \dots, j-1\})$ for every j . Applying twice the inequality of von Bahr and Esseen [20, Theorem 2] yields

$$\mathbb{E}|M_n|^p \leq 2 \sum_{j=2}^n \mathbb{E} \left| \sum_{i=1}^{j-1} f_{ij}(X_i, X_j) \right|^p \leq 2 \sum_{j=2}^n 2 \sum_{i=1}^{j-1} \mathbb{E}|f_{ij}(X_i, X_j)|^p$$

□

2.3. *Proof of Theorem 1.1.* Lemma 2.1 combined with the elementary inequality $|\max a_m - \max b_m| \leq \max |b_m - a_m|$ yields

$$(23) \quad \begin{aligned} & \max_{m=n}^{n+k} T_m \\ &= T_n + \frac{1}{\sqrt{n}} \max_{m=n}^{n+k} \sum_{i=n+1}^m f_0(X_i) + \sqrt{\frac{k}{n}} \varepsilon_{\mathbb{P}}(\sqrt{\frac{k}{n}}) \quad \left(\frac{k}{n} \rightarrow 0\right) \end{aligned}$$

Let $\varepsilon \in]0, 1]$. In this proof, $c_1, c_2, c_3 \in]0, \infty[$ and implied constants in $O(\dots)$ -statements do not depend on n, k, ε , but may depend on α , the law of X_1 , and the sequence (T_n) . Using first (23) and then the independence and sta-

tionarity of the sequence (X_i) , we get

$$\begin{aligned}
(24) \quad & 1 - \alpha_{n,k} \\
&= \mathbb{P}(\max_{m=n}^{n+k} T_m \leq 0) \\
&\leq \mathbb{P}\left(T_n \leq -\frac{1}{\sqrt{n}} \max_{m=n}^{n+k} \sum_{i=n+1}^m f_0(X_i) \pm \varepsilon \sqrt{\frac{k}{n}}\right) + o\left(\sqrt{\frac{k}{n}}\right) \\
&\geq \int_{\mathbb{R}} \mathbb{P}(T_n \leq y) dQ_{n,k}(y) + o\left(\sqrt{\frac{k}{n}}\right)
\end{aligned}$$

for $k/n \rightarrow 0$, with $Q_{n,k}$ denoting the law of

$$Y_{n,k} := -\frac{1}{\sqrt{n}} \max_{\ell=0}^k S_\ell \pm \varepsilon \sqrt{\frac{k}{n}} \quad (n, k \in \mathbb{N})$$

where

$$S_\ell := \sum_{i=1}^{\ell} f_0(X_i) \quad (\ell \in \mathbb{N}_0)$$

Since the $f_0(X_i)$ are i.i.d., we can apply a result of Kac, see [9, Theorem 4.1] and also [19, p. 330], to get

$$(25) \quad \mathbb{E}Y_{n,k} = -\frac{1}{\sqrt{n}} \sum_{\ell=1}^k \frac{1}{\ell} \mathbb{E}(S_\ell)_+ \pm \varepsilon \sqrt{\frac{k}{n}}$$

Since $-Y_{n,k} = \max_{\ell=0}^k (\frac{1}{\sqrt{n}} S_\ell \mp \varepsilon \sqrt{k/n})$ is the maximum of a martingale, the usual L^2 maximum inequality, see e.g. [6, p. 248], yields

$$(26) \quad \mathbb{E}Y_{n,k}^2 \leq 4 \mathbb{E}\left(\frac{1}{\sqrt{n}} S_k \mp \varepsilon \sqrt{\frac{k}{n}}\right)^2 = 4(1 + \varepsilon^2) \frac{k}{n} \leq c_1 \frac{k}{n}$$

and hence in particular, by the Chebyshev and Lyapunov inequalities,

$$(27) \quad \mathbb{P}(|Y_{n,k}| \geq 1) \leq c_1 \frac{k}{n}, \quad |\mathbb{E}Y_{n,k}| \leq \sqrt{c_1} \sqrt{\frac{k}{n}}$$

An application of [15, p. 343, Corollary 10.3.8, with $S_n = \mu(P) + T_n/\sqrt{n}$, P the law of X_1 , $\beta = 1 - \alpha$, $\mu(P) = 0$, $\sigma(P) = 1$, $N_\beta = \Phi^{-1}(1 - \alpha)$, $g_1 = \bar{g}_1 = 0$, $B_0(g_1) = B_0(0) = 0$, $P(f_0(\cdot, P)g_1) = 0$, and with U_n in place of V_n according to Discussion 2.1] yields

$$\begin{aligned}
(28) \quad & \mathbb{P}(T_n \leq y) \\
&= F_n(y) + o\left(\frac{1}{\sqrt{n}}\right) \quad (n \rightarrow \infty, \text{ locally uniformly in } y \in \mathbb{R})
\end{aligned}$$

with

$$(29) \quad F_n(y) := \Phi\left(\Phi^{-1}(1 - \alpha) + y + \frac{ay + by^2}{\sqrt{n}}\right) \quad (n \in \mathbb{N}, y \in \mathbb{R})$$

where $a, b \in \mathbb{R}$ depend only on P, f_0, f_1, f_2 .

Since the functions $\mathbb{P}(T_n \leq \cdot)$ and F_n are $[0, 1]$ -valued, we get

$$\begin{aligned} (30) \quad & \left| \int_{\mathbb{R}} \mathbb{P}(T_n \leq y) \, dQ_{n,k}(y) - \int_{\mathbb{R}} F_n(y) \, dQ_{n,k}(y) \right| \\ & \leq \int_{[-1,1]} |\mathbb{P}(T_n \leq y) - F_n(y)| \, dQ_{n,k}(y) + Q_{n,k}(\mathbb{R} \setminus [-1, 1]) \\ & = o\left(\frac{1}{\sqrt{n}}\right) + O\left(\frac{k}{n}\right) \quad (n \rightarrow \infty) \quad [\text{by (28) and (27)}] \\ & = o\left(\sqrt{\frac{k}{n}}\right) \quad \left(\frac{k}{n} \rightarrow 0\right) \end{aligned}$$

Using

$$F_n(0) = 1 - \alpha, \quad F'_n(0) = \varphi(\Phi^{-1}(1 - \alpha)) \cdot \left(1 + \frac{a}{\sqrt{n}}\right)$$

$$\|F''_n\|_{\infty} := \sup_{y \in \mathbb{R}} |F''_n(y)| \leq c_2 \quad (n \in \mathbb{N})$$

with c_2 depending only on a and b , a Taylor expansion of F_n around zero yields

$$\begin{aligned} (31) \quad & \int_{\mathbb{R}} F_n(y) \, dQ_{n,k}(y) \\ & = F_n(0) + F'_n(0) \mathbb{E}Y_{n,k} + O(\|F''_n\|_{\infty} \mathbb{E}Y_{n,k}^2) \\ & = 1 - \alpha + \varphi(\Phi^{-1}(1 - \alpha)) \left(1 + \frac{a}{\sqrt{n}}\right) \mathbb{E}Y_{n,k} + O\left(\frac{k}{n}\right) \\ & = 1 - \alpha - \frac{\varphi(\Phi^{-1}(1 - \alpha))}{\sqrt{n}} \sum_{\ell=1}^k \frac{1}{\ell} \mathbb{E}(S_{\ell})_+ \pm c_3 \varepsilon \sqrt{\frac{k}{n}} + O\left(\frac{k}{n}\right) \end{aligned}$$

using (26) and (25). Combining (2), (24), (30) and (31) yields (8).

We have $\mathbb{E}(S_{\ell}/\sqrt{\ell})_+ \rightarrow 1/\sqrt{2\pi}$ for $\ell \rightarrow \infty$, by the uniform integrability of $(S_{\ell}/\sqrt{\ell})_+$ following from $\mathbb{E}(S_{\ell}/\sqrt{\ell})_+^2 \leq 1$ and by the central limit theorem, compare [3, Theorems 25.12 and 27.1]. It follows that $\sum_{\ell=1}^k \frac{1}{\ell} \mathbb{E}(S_{\ell})_+ \sim \frac{1}{\sqrt{2\pi}} \sum_{\ell=1}^k \frac{1}{\sqrt{\ell}} \sim \frac{2}{\sqrt{2\pi}} \sqrt{k}$ for $k \rightarrow \infty$. Hence (8) yields (9). \square

2.4. *Proof of Lemma 1.1.* We have

$$(32)h = \frac{1}{\sqrt{2\pi}} \cdot \left(]0, \infty[\ni x \mapsto \frac{1}{x} \right) \circ \frac{1-\Phi}{\varphi} \circ \left(]0, 1[\ni \alpha \mapsto \Phi^{-1}(1-\alpha) \right)$$

Since Mills' ratio $\mathbb{R} \ni x \mapsto \frac{1-\Phi}{\varphi}(x) = \int_0^\infty \exp(-xt - t^2/2) dt$ and the other two composition factors in (32) are strictly decreasing, so is h . Applying the well-known asymptotics $\frac{1-\Phi}{\varphi}(x) \sim \frac{1}{x}$ for $x \rightarrow \infty$ and $\Phi^{-1}(1-\alpha) \sim \sqrt{2 \log(\frac{1}{\alpha})}$ for $\alpha \rightarrow 0$ to (32), we get (10). \square

2.5. *Proof of Lemma 1.2.* By assumption $\lim_{y \rightarrow \infty} \sup_{n \geq n_0} \mathbb{E}Y_n(Y_n > y) = 0$, and for $t \geq 1$

$$\begin{aligned} t\sqrt{n} \mathbb{P}\left(|R_n| > \frac{t\varepsilon}{\sqrt{n}}\right) &\leq \varepsilon^{-p} t^{1-p} n^{\frac{1+p}{2}} \mathbb{E}\left(|R_n|^p \left(|R_n| > \frac{t\varepsilon}{\sqrt{n}}\right)\right) \\ &\leq \varepsilon^{-p} \mathbb{E}\left(Y_n(Y_n > \varepsilon^p \sqrt{n})\right) \end{aligned}$$

\square

Acknowledgements. Thanks to Peter Hall for some helpful comments on a previous version of this paper.

References.

- [1] BHATTACHARYA, R.N. and RANGA RAO, R. (1986). *Normal Approximation and Asymptotic Expansions*. Repr. ed. with corrections & supplemental material. Krieger, Malabar.
- [2] BICKEL, P., GÖTZE, F. and VAN ZWET, W.R. (1986). The Edgeworth expansion for U -statistics of degree two. *Ann. Statist.* **14** 1463–1484.
- [3] BILLINGSLEY, P. (1995) *Probability and Measure*, 3rd ed. Wiley, New York.
- [4] DIACONIS, P. (1978). Statistical problems in ESP research. *Science* **201** 131–136.
- [5] DIACONIS, P. and ZABELL, S. (1991) Closed form summation for classical distributions: variations on a theme of de Moivre. *Statist. Sci.* **6** 284–302.
- [6] DURRETT, R. (2005) *Probability: Theory and Examples*. 3rd Ed. Brooks/Cole, Belmont.
- [7] FELLER, W.K. (1940). Statistical aspects of ESP. *Journal of Parapsychology* **4** 271–298.
- [8] GÖTZE, F. (1987). Approximations for multivariate U -statistics. *J. Multivar. Anal.* **22** 212–229.
- [9] KAC, M. (1954). Toeplitz matrices, translation kernels and a related problem in probability theory. *Duke Math. J.* **21** 501–509.
- [10] KOROLJUK, V.S. and BOROVSKICH, YU.V. (1994). *Theory of U-Statistics*. Kluwer, Dordrecht.
- [11] LEHMANN, E.L. and ROMANO, J.P. (2005). *Testing Statistical Hypotheses*. 3rd Ed. Springer.
- [12] MATTNER, L. (2003). Mean absolute deviations of sample means and minimally concentrated binomials. *Ann. Probab.* **31** 914–925.

- [13] PETROV, V.V. (1995). *Limit Theorems of Probability Theory. Sequences of Independent Random Variables*. Clarendon Press, Oxford.
- [14] PFANZAGL, J. with the assistance of R. HAMBÖKER (1994). *Parametric Statistical Theory*. de Gruyter, Berlin.
- [15] PFANZAGL, J. with the assistance of W. WEFELMEYER (1985). *Asymptotic Expansions for General Statistical Models*. Springer, Berlin.
- [16] PYKE, R. and ROOT, D. (1968). On convergence in r -mean of normalized partial sums. *Ann. Math. Statist.* **39** 379–381.
- [17] ROBBINS, H. (1952). Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc.* **58** 527–535. Reprinted in Robbins (1985), pp. 169–177.
- [18] ROBBINS, H. (1985). *Selected Papers*. Edited by T.L. Lai and D. Siegmund. Springer, N.Y.
- [19] SPITZER, F. (1956). A combinatorial lemma and its application to probability theory. *Trans. Amer. Math. Soc.* **82** 323–339.
- [20] VON BAHR, B. and ESSEEN, C.-G. (1965). Inequalities for the r th absolute moment of a sum of random variables, $1 \leq r \leq 2$. *Ann. Math. Statist.* **36** 299–303.

UNIVERSITÄT TRIER
FACHBEREICH IV - MATHEMATIK
54286 TRIER
GERMANY
E-MAIL: mattner@uni-trier.de